1-1-2012

Water diffusion in and out of the β-barrel of GFP and the fast maturing fluorescent protein, TurboGFP

Marc Zimmer
Connecticut College, mzim@conncoll.edu

Binsen Li
Connecticut College

Ramza Shahid
Connecticut College

Paola Peshkepija
Connecticut College

Recommended Citation

Li, B.; Shahid, R.; Peshkepija, P.; Zimmer, M. Water diffusion in and out of the [Beta symbol]-barrel of GFP and the fast maturing fluorescent protein, turboGFP. *Chem. Phys.* **2012**, *392*, 143-148. http://dx.doi.org/10.1016/j.chemphys.2011.11.001

Published in final edited form as:
Chem Phys. 2012 January 1; 392(1): 143–148. doi:10.1016/j.chemphys.2011.11.001.

Water Diffusion In And Out Of The β-Barrel Of GFP and The Fast Maturing Fluorescent Protein, TurboGFP

Binsen Li, **Ramza Shahid**, **Paola Peshkepija**, and **Marc Zimmer***
Chemistry Department, Connecticut College, New London, CT06320

Abstract

The chromophore of fluorescent proteins is formed by an internal cyclization of the tripeptide 65SYG67 fragment and a subsequent oxidation. The oxidation is slow - the kinetics of this step is presumably improved in fast maturing GFPs. Water molecules can aid in the chromophore formation. We have used 50ns molecular dynamics simulations of the mature and immature forms of avGFP and TurboGFP to examine the diffusion of water molecules in-and-out of the protein β-barrel. Most crystal structures of GFPs have well-structured waters within hydrogen-bonding distance of Glu222 and Arg96. It has been proposed that they have an important role in chromophore formation. Stable waters are found in similar positions in all simulations conducted. The simulations confirm the existence of a pore that leads to the chromophore in the rapidly maturing TurboGFP; decreased water diffusion upon chromophore formation; and increased water diffusion due to the pore formation.

Keywords

Fluorescent Proteins; Chromophore; Chromophore formation; Posttranslational modifications

INTRODUCTION

Green fluorescent protein (GFP) and GFP-like proteins are commonly used in scientific research.(1-6) Their utility as genetic tracer molecules, as highlighters in high resolution microscopy, and as a critical component in many modern biotechnological methods has led to an increased effort to understand their photochemistry, and to find and design new GFP-like proteins. Over 150 distinct GFP-like proteins are currently known. GFP-like fluorescent proteins (FPs) have been found in marine organisms ranging from chordates (e.g. amphioxus) to cnidarians (e.g. corals and sea pansies).(7) Enhanced GFP is still the most commonly used fluorescent protein, despite the fact that it has some major shortcomings such as its slow attainment of fluorescence, especially at 37°C. Second generation fluorescent proteins, such as Vivid Verde (from the *Cyphastrea microphthalma* coral)(8) and TurboGFP (from the *Pontellina plumata* copepoda)(9) overcome these problems.

In *Aequorea victoria* GFP (*av*GFP), the chromophore is formed by an autocatalytic internal cyclization of the tripeptide 65SYG67 fragment and subsequent oxidation of the intrinsically formed structure, see Figure 1. GFP fluorescence is not observed until 90 minutes to 4 hours

*Corresponding author mzim@conncoll.edu (Marc Zimmer) .

after protein synthesis.(10,11) The protein folds quickly and GFP refolding from an acid-, base-, or guanidine HCl-denatured state (chromophore containing but non-fluorescent) occurs with a half-life of between 24 seconds(12) and 5 minutes(13) and the recovered fluorescence is indistinguishable from that of native GFP.(14) The folding of GFP exhibits hysteresis that is due to the decreased flexibility of the chromophore vs. its immature analog.(15) The chromophore formation and oxidation is slow,(16) and it is the kinetics of this step that are presumably improved in fast maturing GFP-like proteins.

The two most probable mechanisms(17) for chromophore formation are a cyclization-oxidation-dehydration mechanism proposed by Wachter(18) and the cyclization-dehydration-oxidation mechanism proposed by Getzoff(19), see Figure 1.

The GFP-maturation kinetics have been separated into pre-oxidation, oxidation and postoxidation events.(20) The preoxidation step, which includes folding and protein backbone condensation, and is thought to lead to intermediate **2** (Figure 1) is the fastest. The following oxidation steps that lead to intermediate **3** (Figure 1) are the slowest. (20) Although the final post-oxidation steps are slow and contribute to rate retardation, they are not the rate limiting steps. Numerous roles have been proposed for water molecules in the postcyclization steps, see Figure 2.

To determine whether the pore in TurboGFP increases water diffusion in and out of the β-barrel, and whether the water molecules might be involved in the chromophore formation we have undertaken 50ns molecular dynamics simulations of avGFP, TurboGFP and the V197L TurboGFP mutant, as well as the immature (precyclized) forms of avGFP and TurboGFP.

MATERIALS AND METHODS

Structural comparisons of the 566 GFP-like chains(23) were done using the protein-ligand database Relibase+ v3.01(24,25). The coordinates of the *Aequorea victoria* GFP crystal structure (1gfl)(26) and *Pontellina plumata* TurboGFP (2g6x)(9) were obtained from the Protein Data Bank (PDB)(27). The protein preparation workflow and Epik v2.0109(28) were used with hydrogen-bond optimization to add hydrogen atoms to protein and solvent atoms as required. The OPLS_2005 force field of MacroModel v9.8107(29) was used.

The starting structure for the immature form of avGFP and TurboGFP, for which no crystal structure has been determined, were calculated by graphically mutating the *Aequorea victoria* GFP (1gfl)(26) and *Pontellina plumata* TurboGFP (2g6x)(9) crystal structures so that the chromophore forming tripeptide sequences were in the original precyclized form and undertaking a conformational search. Conformational searches were conducted using the combined Monte Carlo torsional variation and low mode method.(30,31) The flexible dihedral angles of all the side-chains of residues 64, 65, 66, 67 and 68 (1GFL numbering) were randomly rotated by between 0 and 180° and all solvent molecules in an 8.00 Å sphere from residues 64-68 were randomly rotated and translated by between 0 and 1.00Å in each Monte Carlo (MC) step.(32) 15,000 MC steps were taken in each search. Structures within 50kJ/mol of the lowest energy minimum were kept, and a usage directed method(31) was used to select structures for subsequent MC steps. Structures found in the conformational search were considered unique if the least squared superimposition of equivalent non-hydrogen atoms found one or more pairs separated by 0.25Å or more. The lowest energy structure obtained in the search was further subjected to a 5,000 step large scale low mode conformational search.(33,34) TurboGFP (2g6x) was graphically mutated to get the V197L mutant and conformational searches as described above were used to find low energy conformations.

The final structures obtained from the fully minimized pdb structures and the conformational searches were used to initiate molecular dynamics (MD). MD simulations were carried out in the NPT ensemble at 300K and 1 bar with 1.5fs steps using Desmond(35). All molecular dynamics calculations used the OPLS_2005 force field and SHAKE constrained hydrogens. Ten thousand four hundred and eighteen structures were sampled in each 50ns MD simulation. Each structure was in an orthorhombic simulation box of 0.15M NaCl and SPC waters, with a 10Å solvent buffer between the protein surface and the boundary. Minimizations and pre-equilibrium simulations were done using the default Desmond/multisim relaxation protocol.

Fifty nanosecond simulations of avGFP, TurboGFP and the V197L TurboGFP mutant were conducted with the chromophoric Tyr66 in both the neutral and anionic forms. The immature precyclized forms of avGFP and TurboGFP were only simulated with neutral forms of the chromophoric Tyr66, see Figure 3.

RESULTS AND DISCUSSION

Initially it was thought that one of the functions of the 11-sheet β-barrel structure in fluorescent proteins was to protect the chromophore from its environment, especially from oxygen, which would quench its fluorescence.(36) Recent crystal structures have shown that this does not seem to be the case, as some mutants like YFP have significant holes in the β-barrel resulting in apparent solvent channels from the chromophore to the bulk solvent. Since water molecules may be involved in chromophore formation, proton chains from the chromophore to the surface(37,38) and excited state proton transfer(39) we have examined the pores and the water diffusion in and out of avGFP and TurboGFP.

Based on the crystal structure of TurboGFP it was proposed that a water filled pore leading from the outside to the chromophore may facilitate water and/or oxygen diffusion to the chromophore, thereby speeding up its formation.(9) The pore is lined by residues 136, 137, 156, 197 and 198. In order to confirm that the pore was responsible for the fast chromophore maturation a V197L mutant was created as the increased steric bulk of the leucine would presumably limit the size of the pore. As expected the V197L mutation had markedly slower maturation kinetics, it presumably constricts the pore size.

MolAxis(40,41) was used to determine the channels that connect the chromophore residues to the outside in wild type *Aequorea victoria* GFP (avGFP − chain A in pdb code = 1GFL), TurboGFP and the V197L TurboGFP mutant. Since there is no crystal structure of the V197L mutant available the structure of TurboGFP (pdb code = 2G6X) was graphically mutated to the V197L mutant and a conformational search was undertaken to find the lowest energy structure, which was used in the MolAxis calculations. Figure 4 shows the highest ranked channel for each of the structures, these are the shortest and widest channels. The lower the channel number the higher flux score (shorter and wider). The highest ranked channel for both avGFP and the V197L TurboGFP mutant is a long channel with a constrained bottleneck (0.79 and 1.00Å respectively) that leads from the lids to the chromophore. In TurboGFP the channel with the highest flux score is much shorter and wider, it has a bottleneck radius of 1.23Å and it passes between the beta sheets. The highest ranked channel that passes through the beta sheets in the V197L mutant is channel 5, it has a bottleneck radius of 0.77 Å. In wild type avGFP the highest ranked channel that passes through the beta sheets (Channel 2) has a bottleneck radius of 0.34 Å, see Figure 4.

Both oxygen and water molecules have been implicated in the chromophore formation of GFP-like proteins(17,22,42). We have therefore computationally generated immature precyclized forms of avGFP and TurboGFP from their mature crystallized forms and

subjected them to comprehensive conformational searches. Immature avGFP has a narrow major channel with a bottleneck radius of 0.35 Å, while the precyclized form of TurboGFP's major channel is larger and it goes through the top of the crystal structure and its bottleneck radius is 1.03 Å.

To confirm and further expand on the MolAxis pore analysis a series of fifty nanosecond simulations were undertaken to examine water diffusion in and out of the β-barrel of avGFP and TurboGFP with the chromophore in both its anionic and neutral states, see Figure 3. Identical simulations were undertaken with the V197L mutant of TurboGFP, this mutation was designed to narrow the pore that connects the chromophore cavity to the protein exterior.(9) The mutation significantly slowed the chromophore formation. It has been suggested that the narrower pore restricts oxygen access to the immature chromophore, or impedes the waters that may be involved in proton abstraction from the Y66 Cα atom.(9) Finally two 50ns simulations of avGFP and TurboGFP in their neutral precyclized immature forms were also undertaken.

Table 1 summarizes our findings. As expected more water molecules diffuse in and out of TurboGFP than avGFP. Furthermore the V197L mutation in TurboGFP reduces water diffusion. The FPs with anionic (phenolate) chromophores are more permeable than their neutral counterparts in all the simulations. The autocatalytic cyclization and chromophore formation forms a tighter β-barrel and all simulations of the immature forms have more waters diffusing in and out of the barrel than the mature forms. This can be explained by the fact that the mature fully cyclized FPs are more compact and sterically constrained than the immature uncyclized versions and therefore less porous to water molecules.

All FPs fold into an 11 sheet β-barrel that has the chromophore containing alpha helix running through it. The barrels have two lids, and both the N and C termini compromise part of the same lid. The residues in the lid that contain the N and C termini are much less conserved than those in the lid on the other side(23). In avGFP the waters all enter and leave the β-barrel through a bifurcated pore located on the edge of the more conserved lid, while in TurboGFP most of the flux occurs through the pore adjacent to the chromophore as shown in Figure 4 and first described by Evdokimov et al(9). A small number of molecules enter and exit TurboGFP through temporary pores formed on the edges of the less conserved lids, see Figure 5. The β-barrel of the uncyclized forms of avGFP and TurboGFP are less rigid and temporary pores are formed on the edges of both lids.

Figure 6 shows the path taken by one of the waters escaping from anionic TurboGFP, it diffuse through the pore presumed responsible for the increased maturation rate.(9) Ninety two percent of waters diffuse in and out of this pore in the anionic form of TurboGFP. Analysis of the individual structures collected shows that the pore in TurboGFP is larger than the equivalent pore in avGFP at every stage of the simulation.

Due to the rigid nature of the FP β-barrel most FP crystal structures have numerous well-ordered water molecules, which are well-defined in the FP crystal structures. Intricate hydrogen-bonding networks involving these structural waters possibly have a role in the excited state proton transfer suggested by Agmon.(38) One of these waters is located within hydrogen-bonding distance from Glu222 and in van der Waals contact to the Cα of Tyr66, it is proposed to play a role in chromophore formation, see Figure 2.(17,21) Our simulations have shown that the waters hydrogen bonded to Glu222 in avGFP (and Glu210 in TurboGFP) form strong, stable interactions with the glutamic acid residue. They are one of a few waters that remain in the same position throughout the simulation – in all structures examined. Figure 7 shows the relative positions of the well-structured waters in the crystal

structures of all the GFP-like proteins in the protein databank as compared with their counterparts in the molecular dynamics simulations.

The fact that these waters are found located between the conserved and catalytically important Arg96 and Glu222 residues and the chromophore in both the crystal structures of fluorescent proteins(17) and all our simulations indicates that they are very important and play an active role in chromophore formation.

Examination of the simulations of the precyclized forms of the fluorescent proteins reveal a water molecule that is within hydrogen bonding distance of the amide nitrogens of residues 66 and 67 (avGFP numbering) and the Glu222 sidechain throughout the simulations.

CONCLUSION

From the very first crystal structures of avGFP it was noticed that the β-barrel contained a cavity above the chromophore(44,45). Since then it has been found that the cavity is conserved in all GFP-like proteins and that it contains well-structured water molecules that can be located in most crystal structures. These waters are located between the chromophore and the catalytically active and conserved Arg96 and Glu222 residues. In all our simulations we find that there are waters molecules located in similar positions, they are within hydrogen bonding distance to the Arg96 and Glu222 residues and parts of the chromophore involved in chromophore formation, see Figure 7. Our simulations provide further evidence for the probable role of the structurally conserved water molecules in chromophore formation.

Our simulations show that water diffusion into the β-barrel decreases once the chromophore cyclization has occurred. The more compact mature chromophore allows for a tighter β-barrel than the relaxed precyclized form. This suggests that the β-barrel in GFP-like proteins may have evolved to be porous during chromophore formation, thereby permitting catalytically important water and oxygen molecules to enter, and to be more impermeable following chromophore formation, in this manner protecting the chromophore from quenching. The increased diffusion in and out of the precyclized β-barrels occurs through pores formed between strands 7 and 10, as well as 7 and 8, the very same strands that exhibit increased flexibility in NMR studies.(46,47) This tensing of the β-barrel upon chromophore formation could be the cause for the hysteretic folding behavior observed by Jennings et al. (15,47,48)

The phenolate form of avGFP is more permeable than the neutral phenol form. The enhanced water diffusion in the anionic form is a result of the main pore being more flexible thereby facilitating water movement in and out of the cavity.

TurboGFP is a second generation fluorescent protein with significantly improved chromophore maturation kinetics. Based on the crystal structure of TurboGFP it was proposed that a water filled pore leading from the outside to the chromophore may facilitate water and/or oxygen diffusion to the chromophore, thereby speeding up its formation.(9) Our simulations show that TurboGFP is indeed significantly more permeable to water molecules than GFP, and that most of the water molecules enter through the primary pore located adjacent to the chromophore. The 11-sheet β-barrel plays a determining role in the maturation rate by controlling the number of water molecules diffusing into the can. Due to TurboGFP's unique shorter and wider water-exchanging pore structure, which passes between its β-sheets, more water molecules diffuse into the β-barrel compared to avGFP - this leads to its much more rapid maturation. The V197L mutation in TurboGFP had markedly slower maturation kinetics, it presumably constricts the pore size. This was

confirmed by our simulations, which show that the mutation significantly reduces water diffusion in and out of the β-barrel.

Acknowledgments

This research was supported by an allocation of advanced computing resources provided by the National Science Foundation. The computations were performed on Kraken at the National Institute for Computational Sciences (http://www.nics.tennessee.edu/). MZ is the Jean Tempel '65 Professor of Chemistry. The research was funded by the NIH (Area Grant # R15 GM59108-02). NSF-STEM 0726589 for RS and PP.

REFERENCES

(1). Craggs TD. Green fluorescent protein: structure, folding and chromophore maturation. Chemical Society Reviews. 2009; 38:2865–75. [PubMed: 19771333]

(2). Day RN, Davidson MW. The fluorescent protein palette: tools for cellular imaging. Chemical Society Reviews. 2009; 38:2887–921. [PubMed: 19771335]

(3). Zimmer, M. Glowing Genes: A Revolution in Biotechnology. Prometheus Books; Amherst, N.Y.: 2005.

(4). Zimmer M. GFP: from jellyfish to the Nobel prize and beyond. Chemical Society Reviews. 2009; 38:2823–32. [PubMed: 19771329]

(5). Chudakov DM, Matz MV, Lukyanov S, Lukyanov KA. Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues. Physiological Reviews. 2010; 90:1103–63. [PubMed: 20664080]

(6). Shaner NC, Patterson GH, Davidson MW. Advances in fluorescent protein technology. J Cell Sci. 2007; 120:4247–60. [PubMed: 18057027]

(7). Alieva NO, Konzen KA, Field SF, Meleshkevitch EA, Beltran-Ramirez V, Miller DJ, Salih A, Wiedenmann J, Matz MV. Diversity and Evolution of Coral Fluorescent Proteins. PLOS ONE. 2008; 3:e2680. [PubMed: 18648549]

(8). Ilagan RP, Rhoades E, Gruber DF, Kao HT, Pieribone VA, Regan L. A new bright green-emitting fluorescent protein - engineered monomeric and dimeric forms. Febs Journal. 2010; 277:1967–78. [PubMed: 20345907]

(9). Evdokimov AG, Pokross ME, Egorov NS, Zaraisky AG, Yampolsky IV, Merzlyak EM, Shkoporov AN, Sander I, Lukyanov KA, Chudakov DM. Structural basis for the fast maturation of Arthropoda green fluorescent protein. Embo Reports. 2006; 7:1006–12. [PubMed: 16936637]

(10). Crameri A, Whitehorn EA, Tate E, Stemmer WPC. Improved green fluorescent protein by molecular evolution using DNA shuffling. Nat. Biotechnol. 1996; 14:315–19. [PubMed: 9630892]

(11). Heim R, Prasher DC, Tsien RY. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. Proc. Natl. Acad. Sci. USA. 1994; 91:12501–04. [PubMed: 7809066]

(12). Makino Y, Amada K, Taguchi H, Yoshida M. Chaperon-mediated folding of green fluorescent protein. J. Biol. Chem. 1997; 272:12468–74. [PubMed: 9139695]

(13). Ward WW, Bokman SH. Reversible denaturation of Aequorea Green-Fluorescent Protein: Physical separation and characterization of the renatured protein. Biochemistry. 1982; 21:4535–40. [PubMed: 6128025]

(14). Bokman SH, Ward WW. Renaturation of Aequorea green fluorescent protein. Biochem. Biophys. Res. Commun. 1981; 101:1372–80. [PubMed: 7306136]

(15). Andrews BT, Roy M, Jennings PA. Chromophore Packing Leads to Hysteresis in GFP. Journal of Molecular Biology. 2009; 392:218–27. [PubMed: 19577576]

(16). Reid BG, Flynn GC. Chromophore Formation in Green Fluorescent Protein. Biochemistry. 1997; 36:6786–91. [PubMed: 9184161]

(17). Lemay NP, Morgan AL, Archer EJ, Dickson LA, Megley CM, Zimmer M. The role of the tight-turn, broken hydrogen bonding, Glu222 and Arg96 in the post-translational green fluorescent protein chromophore formation. Chemical Physics. 2008; 348:152–60. [PubMed: 19079566]

(18). Rosenow MA, Huffman HA, Phail ME, Wachter RM. The crystal structure of the Y66L variant of green fluorescent protein supports a cyclization-oxidation-dehydration mechanism for chromophore maturation. Biochemistry. 2004; 43:4464–72. [PubMed: 15078092]

(19). Barondeau DP, Putnam CD, Kassmann CJ, Tainer JA, Getzoff ED. Mechanism and energetics of green fluorescent protein chromophore synthesis revealed by trapped intermediate structures. Proceedings of the National Academy of Sciences of the United States of America. 2003; 100:12111–16. [PubMed: 14523232]

(20). Zhang LP, Patel HN, Lappe JW, Wachter RM. Reaction progress of chromophore biogenesis in green fluorescent protein. Journal of the American Chemical Society. 2006; 128:4766–72. [PubMed: 16594713]

(21). Pouwels LJ, Zhang LP, Chan NH, Dorrestein PC, Wachter RM. Kinetic isotope effect studies on the de novo rate of chromophore formation in fast- and slow-maturing GFP variants. Biochemistry. 2008; 47:10111–22. [PubMed: 18759496]

(22). Wachter RM. Chromogenic cross-link formation in green fluorescent protein. Accounts of Chemical Research. 2007; 40:120–27. [PubMed: 17309193]

(23). Ong WJH, Alvarez S, Leroux IE, Shahid RS, Samma AA, Peshkepija P, Morgan AL, Mulcahy S, Zimmer M. Function and structure of GFP-like proteins in the protein data bank. Molecular BioSystems. 2011; 7:984–92. [PubMed: 21298165]

(24). Hendlich M, Bergner A, Gunther J, Klebe G. Relibase: Design and development of a database for comprehensive analysis of protein-ligand interactions. Journal of Molecular Biology. 2003; 326:607–20. [PubMed: 12559926]

(25). Gunther J, Bergner A, Hendlich M, Klebe G. Utilising structural knowledge in drug design strategies: Applications using relibase. Journal of Molecular Biology. 2003; 326:621–36. [PubMed: 12559927]

(26). Yang F, Moss LG, Phillips GN. The molecular structure of green fluorescent protein. Nature Biotechnology. 1996; 14:1246–51.

(27). Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Databank. Nucleic Acids Research. 2000; 28:235–42. [PubMed: 10592235]

(28). Shelley, JC.; Cholleti, A.; Frye, LL.; Greenwood, JR.; Timlin, MR.; Uchimaya, M. Epik: a software program for pKa prediction and protonation state generation for drug-like molecules. San Francisco, CA: 2006. p. 681-91.

(29). Mohamadi F, Richards NGJ, Guida WC, Liskamp R, Lipton M, Caufield C, Chang G, Hendrickson T, Still WC. Macromodel - an Integrated Software System for Modeling Organic and Bioorganic Molecules Using Molecular Mechanics. Journal of Computational Chemistry. 1990; 11:440–67.

(30). Chang G, Guida WC, Still WC. An internal-coordinate monte carlo method for searching conformational space. J. Am. Chem. Soc. 1989; 111:4379–86.

(31). Saunders M, Houk K, Wu Y-D, Still W, Lipton M, Chang G, Guida W. Conformations of cycloheptadecane. A comparison of methods for conformational searching. J. Am. Chem. Soc. 1990; 112:1419.

(32). Bartol J, Comba P, Melter M, Zimmer M. Conformational searching of transition metal compounds. J Comput Chem. 1999; 20:1549–58.

(33). Kolossvary I, Keseru GM. Hessian-free low-mode conformational search for large-scale protein loop optimization: Application to c-jun N-terminal kinase JNK3. Journal of Computational Chemistry. 2001; 22:21–30.

(34). Keseru GM, Kolossvary I. Fully flexible low-mode docking: Application to induced fit in HIV integrase. Journal of the American Chemical Society. 2001; 123:12708–09. [PubMed: 11741448]

(35). Bowers, KJ.; Chow, E.; Xu, H.; Dror, RO.; Eastwood, MP.; Gregersen, BA.; Klepeis, JL.; Kolossvary, I.; Moraes, MA.; Sacerdoti, FD.; Salmon, JK.; Shan, Y.; Shaw, DE. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters; Proceedings of the 2006 ACM/IEEE Conference on Supercomputing (SC06); 2006; Tampa, FL: ACM Press;

(36). Zimmer M. Green fluorescent protein (GFP): Applications, structure, and related photophysical behavior. Chemical Reviews. 2002; 102:759–81. [PubMed: 11890756]

(37). Agmon N. Proton pathways in green fluorescence protein. Biophysical Journal. 2005; 88:2452–61. [PubMed: 15681647]

(38). Shinobu A, Palm GJ, Schierbeek AJ, Agmon N. Visualizing Proton Antenna in a High-Resolution Green Fluorescent Protein Structure. Journal of the American Chemical Society. 2010; 132:11093–102. [PubMed: 20698675]

(39). Jain RK, Ranganathan R. Local complexity of amino acid interactions in a protein core. Proceedings of the National Academy of Sciences of the United States of America. 2004; 101:111–16. [PubMed: 14684834]

(40). Yaffe E, Fishelovitch D, Wolfson HJ, Halperin D, Nussinov R. MolAxis: Efficient and accurate identification of channels in macromolecules. Proteins-Structure Function and Bioinformatics. 2008; 73:72–86.

(41). Yaffe E, Fishelovitch D, Wolfson HJ, Halperin D, Nussinov R. MolAxis: a server for identification of channels in macromolecules. Nucleic Acids Research. 2008; 36:W210–W15. [PubMed: 18448468]

(42). Barondeau DP, Tainer JA, Getzoff ED. Structural evidence for an enolate intermediate in GFP fluorophore biosynthesis. Journal of the American Chemical Society. 2006; 128:3166–68. [PubMed: 16522096]

(43). Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. Journal of Molecular Graphics. 1996; 14:33. &. [PubMed: 8744570]

(44). Ormoe M, Cubitt AB, Kallio K, Gross LA, Tsien RY, Remington SJ. Crystal structure of the Aequorea victoria green fluorescent Protein. Science. 1996; 273:1392–95. [PubMed: 8703075]

(45). Yang F, Moss L, Phillips G. The molecular structure of green fluorescent protein. Nature Biotech. 1996; 14:1246–51.

(46). Seifert MHJ, Georgescu J, Ksiazek D, Smialowski P, Rehm T, Steipe B, Holak TA. Backbone dynamics of green fluorescent protein and the effect of histidine 148 substitution. Biochemistry. 2003; 42:2500–12. [PubMed: 12614144]

(47). Andrews BT, Gosavi S, Finke JM, Onuchic JN, Jennings PA. The dual-basin landscape in GFP folding. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105:12283–88. [PubMed: 18713871]

(48). Andrews BT, Schoenfish AR, Roy M, Waldo G, Jennings PA. The rough energy landscape of superfolder GFP is linked to the chromophore. Journal of Molecular Biology. 2007; 373:476–90. [PubMed: 17822714]

Highlights

Simulations confirm

- decreased water diffusion upon chromophore formation.
- increased water diffusion due to the pore formation, therefore fast chromophore maturation.
- possible role of the structurally conserved water molecules in chromophore formation.

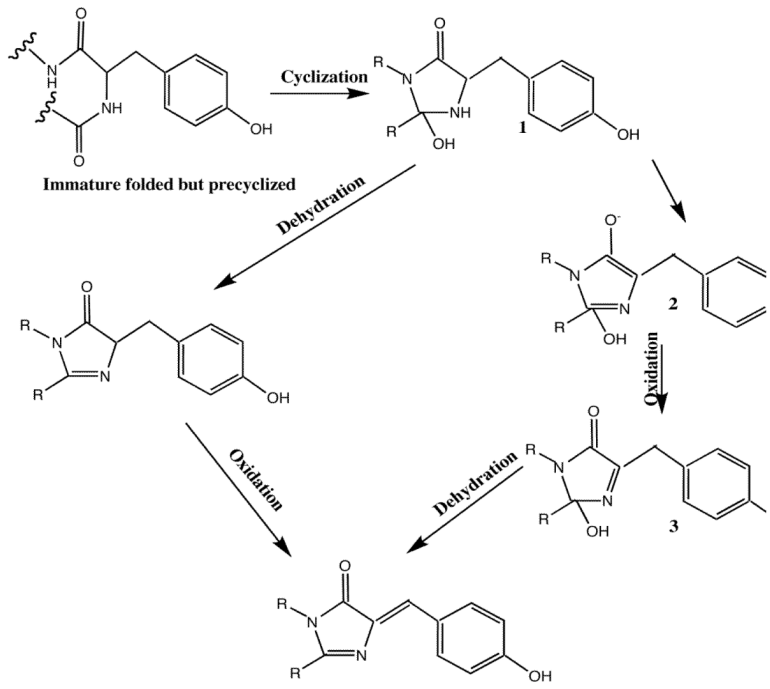

Figure 1.
The cyclization-oxidation-dehydration mechanism proposed by Wachter(18) (right) and the cyclization-dehydration-oxidation mechanism proposed by Getzoff(19) (left).

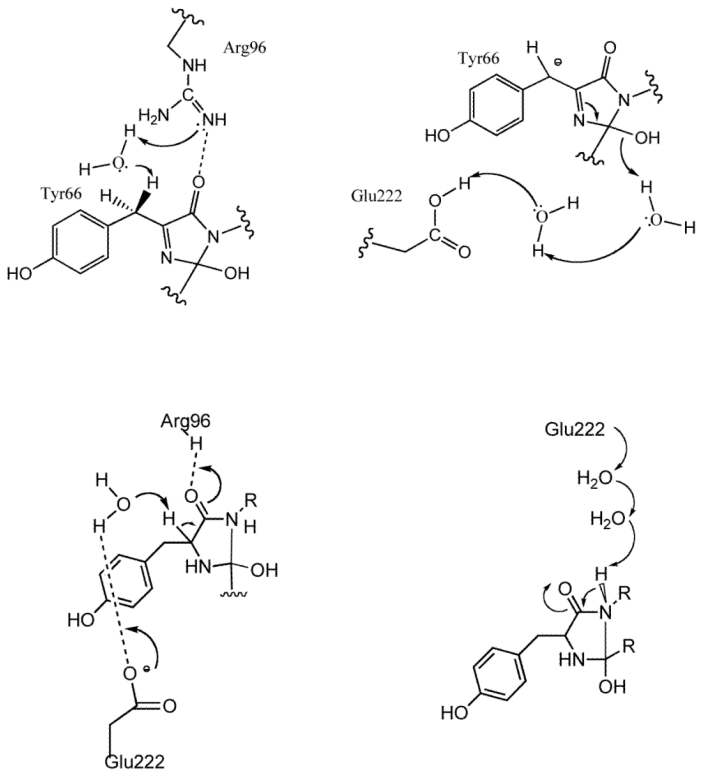

Figure 2.
Possible roles played by the structurally conserved waters in the chromophore formation(1,21). The cyclization-oxidation-dehydration mechanism proposed by Wachter(18,22) (top and bottom left) and the cyclization-dehydration-oxidation mechanism proposed by Getzoff(19) (bottom right).

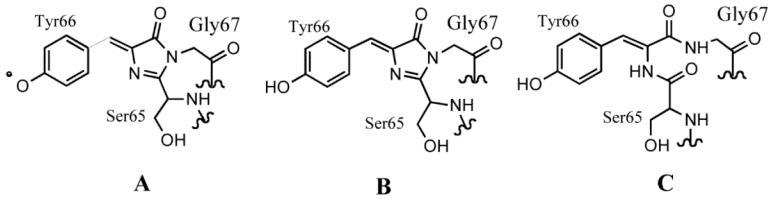

Figure 3.
Anionic (A) and neutral (B) forms of the chromophore. Neutral immature form (C).

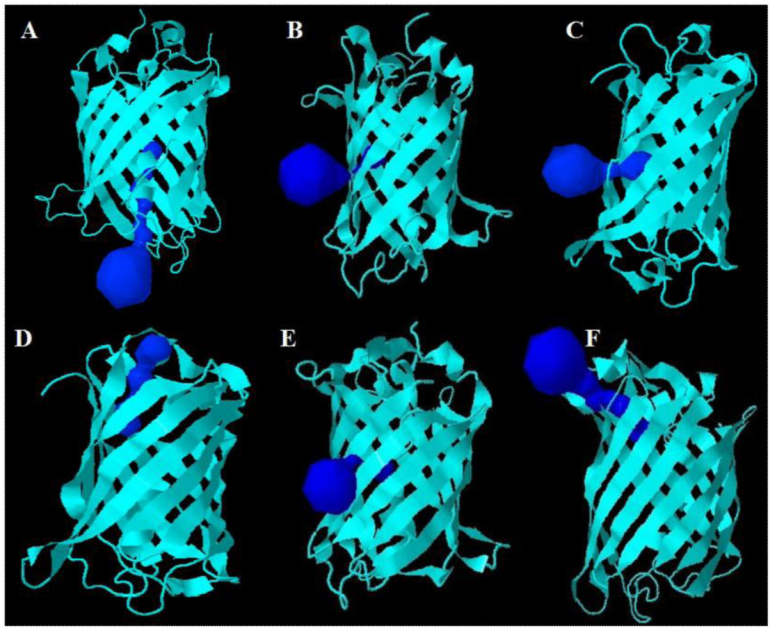

Figure 4.

Highest ranked channel MolAxis(40,41) pores and channels for each of the structures. The lower the channel number the higher flux score (shorter and wider). **A**) avGFP channel 1: Bottleneck radius 0.79 Å **B**) avGFP channel 2: Bottleneck radius 0.34 Å **C**) TurboGFP channel 1: Bottleneck radius 1.23 Å. (Note how channel 2 in avGFP and channel 1 in TurboGFP exit at the same point, however avGFP has a much smaller bottleneck.) **D**) TurboGFP V197L mutant channel 1: Bottleneck radius 1.00 Å **E**) Immature 1GFL channel 1: Bottleneck radius 0.35 Å **F**) immature TurboGFP channel 1: Bottleneck radius 1.03 Å.

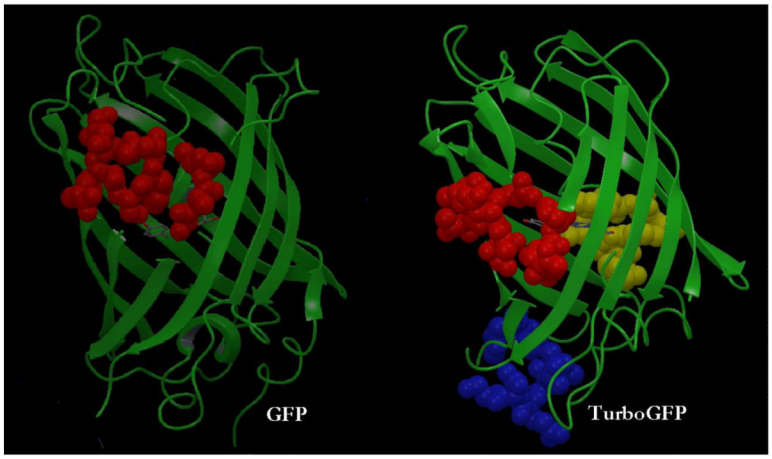

Figure 5.

Residues surrounding bifurcated pore in avGFP are indicated by red CPK spheres (left). The residues surrounding the pore purportedly responsible for TurboGFP's rapid chromophore formation are shown as red CPK spheres (right). The chromophore, tube representation, is visible through the pore. The pores surrounded by the residues colored blue and yellow are located on the lid-edges and are less permanent.

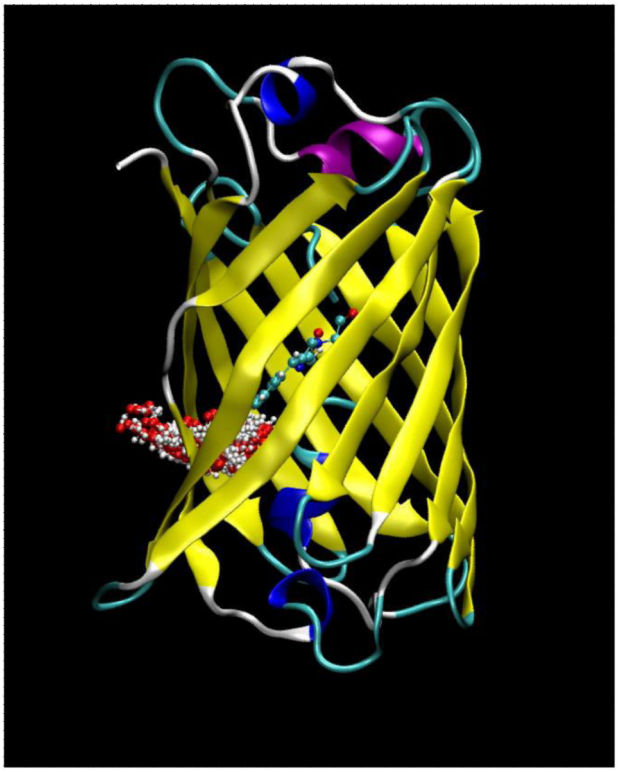

Figure 6.
700 snapshots taken during the 3.6ns it takes one of the water molecules (red oxygen and white hydrogen) to exit through the primary pore of the TurboGFP β-barrel. Water positions were superimposed on the average TurboGFP structure using VMD. (43)

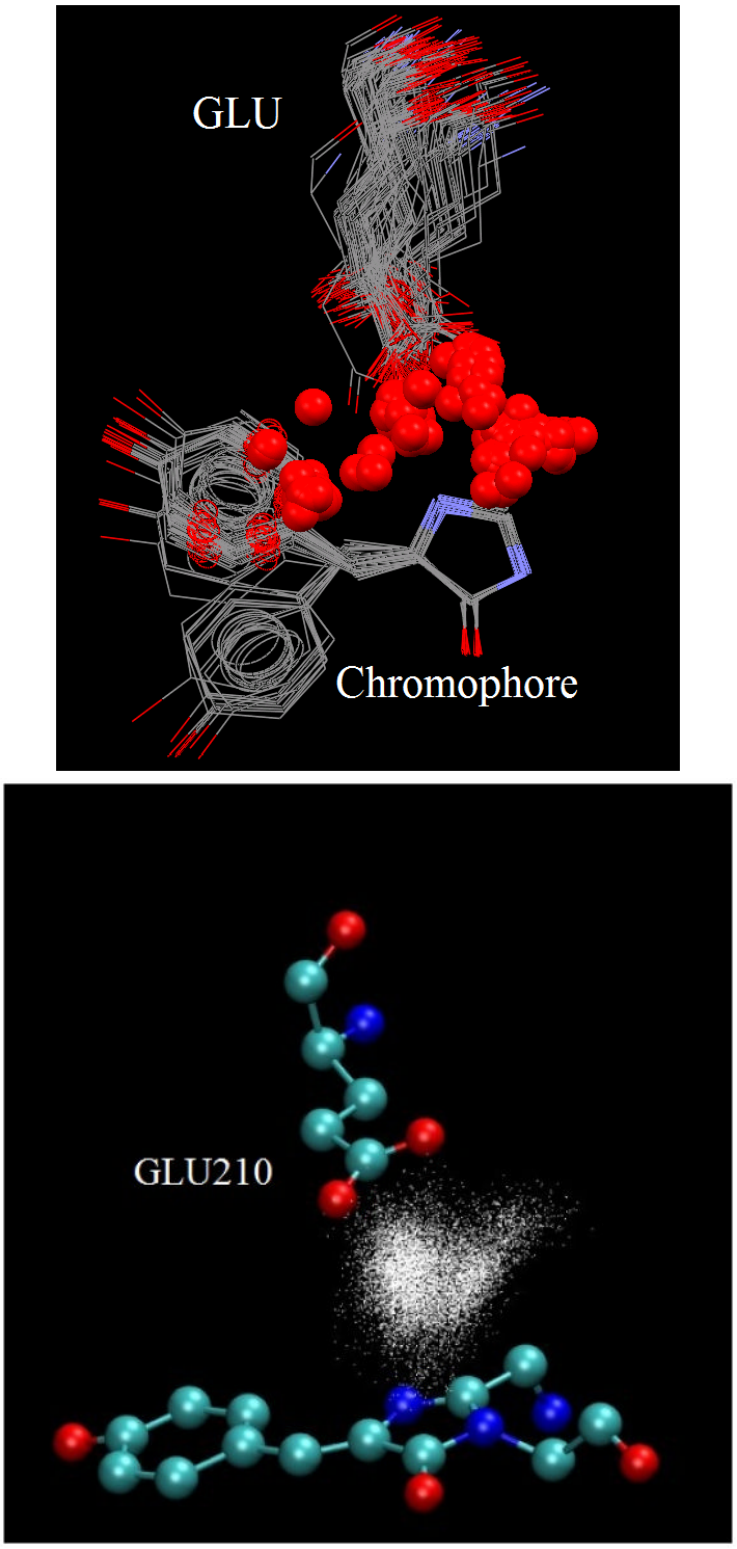

Figure 7.

Left: An Isostar overlay plot of all the crystal structures of GFP and GFP-like proteins with fully cyclized chromophores found with Relibase+ in the Protein Databank. The

imidazolone rings of all the structures were overlapped in order to show the orientation of the glutamic acid residues (Glu222 in GFP numbering) and the water molecule discussed in the text relative to the position of the chromophoric imidazolone ring. **Right:** 10,416 snapshots taken during the 50 ns simulation of the anionic form of TurboGFP. Water positions (white dots) were superimposed on the average TurboGFP structure using VMD. Only the chromophore and Glu210 of average TurboGFP shown. (43)

Table 1

Number of waters diffusing in and out of the fluorescent protein β-barrel and the number of pores involved.

Protein (Chromophore protonation state)	# waters diffusing in and out of β-barrel	Entrance/exit points
GFP (Tyr-OH)	3	1
GFP (Tyr-O^-)	20	1
TurboGFP (Tyr-OH)	12	1
TurboGFP (Tyr-O^-)	50	3
TurboV197LGFP (Tyr-OH)	9	2
TurboV197LGFP (Tyr-O^-)	11	1
Immature GFP (Tyr-OH)	23	3
Immature TurboGFP (Tyr-OH)	46	1